Claims

1. A method of producing a soluble bioactive domain of a protein of interest, the method comprising the step of selecting at least one candidate soluble domain of the protein and assessing the produced protein of each domain for desired activity.

2. The method according to claim 1 comprising the step of amplifying DNA encoding at least one candidate soluble domain, cloning the amplified DNA encoding each candidate domain into at least one expression vector, using each of said vectors into which the DNA has been cloned to each transfect or transform one or more host cell strains, expressing said DNA in one or more of said host cell strains, and analysing expression products from said host cells for solubility.

3. The method according to claim 2 comprising steps:

(a) analysing DNA coding for the protein of interest to identify one or more candidate soluble domains

(b) providing oligonucleotide primers to amplify DNA encoding each domain

(c) amplifying said DNA with said primers

(d) cloning amplified DNA from step (c) for each domain into at least one expression vector

(e) optionally screening clones for correct orientation of DNA

(f) using each of the vectors of step (d) into which the DNA has been cloned to each transfect or transform one or more host cell strains,

(g) expressing said DNA in one or more of said host cell strains, and

(h) analysing expression products from said host cells for solubility.

4. The method according to claim 2 or claim 3 comprising the step of producing a soluble bioactive protein domain of said protein of interest.

5. The method according to any one of claims 2 to 4 wherein at least three candidate soluble domains are selected and DNA is amplified for each of said domains.

6. The method according to any one of claims 2 to 5 wherein said DNA encoding each selected domain is amplified under at least two, preferably at least three different PCR programs in parallel.

7. The method according to claim 6 wherein said PCR programs are selected from (i) a standard PCR programme using a predicted annealing temperature for the primers; (ii) a standard PCR programme using a temperature in the range 48 to 52°C, preferably 50°C as the temperature

for annealing and (iii) a touchdown PCR programme, where the annealing temperature starts at a temperature in the range 62 to 67°C, preferably 65°C, and then gradually decreases to a temperature in the range 48 to 52°C, preferably 50°C, over the subsequent cycles.

8. The method according to any one of claims 2 to 7 wherein the amplified DNA encoding each domain is cloned into a plurality of different expression vectors.

9. The method according to claim 8 wherein the plurality of vectors include one or more of a vector capable of encoding a fusion protein with a poly-Histidine tag, a vector capable of conferring tight regulation of translation to impose stringent expression conditions, a vector capable of encoding a fusion protein with a solubility enhancing tag.

10. The method according to claim 9 wherein the solubility enhancing tag comprises a glutathione-S-transferase tag, a dihydrofolate reductase tag, a NusA tag or a SNUT tag.

11. The method according to any one of claims 2 to 10 wherein the vectors are each transfected or transformed into a plurality of different host cell strains

12. The method according to any one of claims 2 to 11 wherein the host cell strains are different *E. coli* strains.

13. The method according to claim 12 wherein the E coli strains are selected from Rosetta(DE3)pLacI, Tuner(DE3)pLacI, Origami BL21(DE3)pLacI and TOP10F'.

14. The method according to any one of claims 2 to 13 including the step of screening transformants for correct orientation of DNA.

15. The method according to claim 14 wherein the step of screening transformants for correct orientation of the insert is performed using dot-blotting.

16. The method according to any one of claims 2 to 14 wherein the expression products from said host cells are analysed using ELISA or dot-blotting methods.

17. The method according to any one of the preceding claims wherein analysis of expression products includes the use of chloroform and UV light to stain protein on an SDS-PAGE gel.

18. The method according to claim 17, wherein the method further comprises the subsequent use of the chloroform-stained SDS-PAGE gel for western blotting for the identification of proteins.

19. The method according to any one of the preceding claims wherein the protein of interest is a protein encoded by the *yotiao* gene, the murine MAR1 protein or the human Jak1 protein.

20. A method of producing a soluble bioactive domain of a protein of interest comprising the steps:
(a) analysing DNA coding for the protein of interest to identify one or more candidate soluble domains
(b) providing oligonucleotide primers to amplify DNA encoding each domain
(c) amplifying said DNA using, in parallel, a standard PCR programme using a predicted annealing temperature for the primers; (ii) a standard PCR programme using a temperature in the range 48 to 52°C, preferably 50°C, as the temperature for annealing and (iii) a touchdown PCR programme, where the annealing temperature starts at a temperature in the range 62 to 67°C, preferably 65°C, and then gradually decreases to a temperature in the range 48 to 52°C, preferably 50°C, over the subsequent cycles.
(d) cloning amplified DNA from step (b) into a plurality of different expression vectors,
(e) optionally screening clones for correct orientation of DNA
(f) using each of the vectors of step (d) into which the DNA has been cloned to each transfect

or transform a plurality of different host cell strains

(g) expressing said DNA in one or more of said host cell strains, and

(h) analysing expression products from said host cells for solubility.

21. The method according to claim 20 wherein at least three candidate soluble domains are selected and DNA is amplified for each of said domains.

22. The method according to claim 20 or claim 21 wherein the plurality of vectors include one or more of a vector capable of encoding a fusion protein with a poly-Histidine tag, a vector capable of conferring tight regulation of translation to impose stringent expression conditions, a vector capable of encoding a fusion protein with a solubility enhancing tag.

23. The method according to claim 22 wherein the solubility enhancing tag comprises a glutathione-S-transferase tag, a dihydrofolate reductase tag, a NusA tag or a SNUT tag.

24. The method according to any one of claims 20 to 23 wherein the host cell strains are different *E. coli* strains.

25. The method according to claim 24 wherein the E coli strains are selected from

Rosetta(DE3)pLacI, Tuner(DE3)pLacI, Origami B21(DE3)pLacI and TOP10F.

26. A soluble bioactive domain of a protein produced by the method according to any one of claims 1 to 25.

27. Use of a sortase gene product as a purification tag.

28. The use according to claim 27 wherein the sortase gene product is a Staphylococcus aureus srtA gene product.

29. The use according to claim 27 or claim 28 wherein the sortase gene product is encoded by the nucleotide sequence shown in Figure 8 or a variant or fragment thereof.

30. The use according to any one of claims 27 to 29 wherein the sortase gene product comprises amino acids 26 to 171 of the SrtA sequence shown in Figure 8 or a variant or fragment thereof.

31. An expression construct for the production of recombinant polypeptides, which construct comprises an expression cassette consisting of the following elements that are operably linked: a) a promoter; b) the coding region of a DNA encoding a sortase gene product as a purification tag sequence; and c) a cloning

site for receiving the coding region for the recombinant polypeptide to be produced; and d) transcription termination signals.

32. The expression construct according to claim 31 wherein the sortase gene product is a Staphylococcus aureus srtA gene product.

33. The expression construct according to claim 31 or claim 32 wherein the sortase gene product is encoded by the nucleotide sequence shown in Figure 8 or a variant or fragment thereof.

34. The expression construct according to any one of claims 31 to 33 wherein the sortase gene product comprises amino acids 26 to 171 of the SrtA sequence shown in Figure 8 or a variant or fragment thereof.

35. A method for producing a polypeptide, comprising: a) preparing an expression vector for the polypeptide to be produced by cloning the coding sequence for the polypeptide into the cloning site of an expression construct as claimed in any one of claims 30 to 34; b) transforming a suitable host cell with the expression construct thus obtained; and c) culturing the host cell under conditions allowing expression of a fusion polypeptide consisting of the amino acid sequence of the purification tag with the amino acid sequence of the polypeptide to be expressed covalently

linked thereto; and d) isolating the fusion polypeptide from the host cell or the culture medium by means of binding the fusion polypeptide present therein through the amino acid sequence of the purification tag.

36. The method according to claim 35, wherein the sortase gene product is a Staphylococcus aureus srtA gene product.

37. The method according to claim 35 or claim 36 wherein the sortase gene product is encoded by the nucleotide sequence shown in Figure 8 or a variant or fragment thereof.

38. The method according to any one of claims 37 to 35 wherein the sortase gene product comprises amino acids 26 to 171 of the SrtA sequence shown in Figure 8 or a variant or fragment thereof.

39. A fusion polypeptide obtained by the method of any one of claims 35 to 38.

40. A purification tag comprising a sortase gene product.

41. The purification tag according to claim 40 wherein the gene product is a Staphylococcus aureus srtA gene product.

42. The purification tag according to claim 40 or claim 41 wherein the sortase gene product is encoded by the nucleotide sequence shown in Figure 8 or a variant or fragment thereof.

43. The purification tag according to any one of claims 40 to 42 wherein the sortase gene product comprises amino acids 26 to 171 of the SrtA sequence shown in Figure 8 or a variant or fragment thereof.